

JC12 Rec'd PCT/PTC 18 APR 2005

METHOD TO EVALUATE WHETHER A TIME DELAY IS BETTER THAN A LIMIT

5

The invention concerns a method to evaluate whether a time delay as an implementation dependent parameter is better than a statistically defined soft time limit. The invention is especially applied to a class of measurements measuring the delay time from a cell quality swap, generated by a system simulator, to the registration message, generated by the user equipment, a mobile station for example. In a mobile communication system the mobile station (user equipment) should make a cell reselection or handover to another base station of another communication cell if the quality of communication with the current base station of the current cell (cell quality) decreases and the communication quality with an other base station of an other cell increases over the quality of the current base station.

Such a soft handover handled by a mobile station in a communication system with several base stations for a mobile system of the third generation using Code Division Multiple Access (CDMA) is known from US 5,267,261 for example.

The communication standard defines a maximum delay time from the swap cell quality until the time where the user equipment issues a registration message in order to register to the other base station. However, this time delay TD is not defined as a hard limit, i. e. the user equipment would not fail the test if the delay time exceeds the time limit only a single time, but is defined as a soft time limit, i. e. the user equipment shall fulfil the test requirement for a certain percentage (for example 90%) of the cases in repeated measurements. The pass fail decision of the test-equipment shall be done

with a certain quality, for example 3% wrong decision risk.

From the present state of the art it is not known how to
5 deal with such statistically defined soft time limits for repeated tests.

Such a test is time consuming. This is illustrated by the following example: The success ratio SR in this example is
10 90%. The time limit in this example is 8s. Due to statistical significance it is not enough to observe 9 successes in 10 measurements. For statistical significance it is usual to observe $200 \cdot 9$ successes in $200 \cdot 10$ measurements. This test lasts 4.5 hours.

15 It is the object of the present invention to provide an effective method to make a pass fail decision, especially on repeated time delay measurements, against a statistically defined soft time limit.

20 It is further the object of the present invention to provide an effective method to make the pass fail decision in the minimum possible test time preserving the statistical significance.

25 The object is solved alternatively by the features of claim 1, 2, 7 or 8.

The depended claims concern further developments of the
30 invention. Claims 20 to 23 cover the respective digital storage mediums computer program and computer program product.

The invention is further described with respect to the
35 drawings. In the drawings

Fig. 1 shows the scenario of cell swap and measurement of delay time;

- Fig. 2 shows a block diagram of the measurement set up;
- Fig. 3 shows a diagram to illustrate the inventive method,
- 5 Fig. 4 shows the referenced Error Ratio er_{norm} as a function of the measured errors ne ,
- 10 Fig. 5 shows a diagram to illustrate a measurement using a first embodiment of the inventive method,
- 15 Fig. 6 shows a diagram to illustrate a measurement using a second embodiment of the inventive method and
- 20 Fig. 7 shows a diagram illustrating the position at the end of the test using the first embodiment of the inventive method as a function of probability.
- 25 Fig. 1 shows the scenario of the swap and measurement of time delay. Fig. 2 shows the respective measurement set up.
- 30 A system simulator SS simulates the two cells cell 1 and cell 2 of a cellular mobile communication system. A User Equipment UE (mobile station) is registered to one of the cells of the cellular mobile communication system. The system simulator SS swaps the cell quality at specific cyclic time points T1 and T2. From time T1 until time T2 cell 1 has the better cell quality compared to cell 2. For example the field strength received by the User Equipment UE from cell 1 is 3dB higher than received from cell 2.
- 35 From T2 to T1 the cell quality of cell 2 is better as the cell quality of cell 1. The User Equipment UE should recognise the cell swap CS and should react to the cell swap CS by issuing a registration message RM. With the registration message RM the User Equipment UE registers to

the cell serving with better cell quality. The User Equipment UE needs a specific time delay TD to evaluate and recognise the cell swap CS and to react thereon with the registration message RM. This time delay TD is measured for a specific number of tests by the system simulator SS and the result is transferred to the controller CO, as shown in Fig. 2. According to the inventive method the time delays TD are measured. Each time delay TD is hard decided, if it is worse or better than the time limit. Then it is evaluated whether the delays are better than the time-limit in 90% of the cases. After sufficient repetitions an overall pass fail decision with a wrong decision risk of e.g. 3% is done in a minimum possible test time.

15

The following realisation is exploited:

The occurrence of the bad delay in all delays, measured, is statistically independent, or can easily be made statistically independent, by inserting sufficient recovery time after a registration event. Statistically independent events are determined by the binomial distribution for any error ratio. (Error Ratio $ER = 1 - SR$ Success Ratio)

For small error ratios the binomial distribution is approximated by the Poisson distribution. The use of both distributions give meaningful results: The Poisson distribution give optimal results in terms of test time and statistical significance for small error ratios, and the Poisson distribution is well supported in mathematical software tools.

For high error ratios the Poisson distribution give sub-optimal results. The results are conservative in the sense: Test time is longer (worse) than optimal. Wrong decision risk is lower (better) than predefined.

The binomial distribution is optimal in terms of test time and statistical significance for any error ratio. However

the binomial distribution is not so well supported in standard mathematical software tools. Both derivations are described in the following.

5 According to the present invention, the worst possible likelihood distribution PD_{high} and/or the best possible likelihood distribution PD_{low} are obtained by the formulas given in claim 1 or 2 and 7 or 8, respectively. From these likelihood distributions, the average number NE_{high} or
10 NE_{low} is obtained and compared with NE_{limit} , respectively. If NE_{limit} is higher than the average number NE_{high} or smaller than the average number NE_{low} , the test is stopped and it is decided that the device has early passed or early failed, respectively. During the test a single step
15 wrong decision probability for a preliminary ER stage can be used, which is smaller than the probability for the entire test.

The Poisson distribution preferably used for the present
20 invention, is best adapted to a small number of "bad delay times". A "bad delay time" is defined as a delay time, exceeding the time limit. For example a measured time delay exceeding the time limit of 8s would be a "bad time delay", also referred to as an "error". A measured time
25 delay TD below the time limit, of for example 8s, would be a "good time delay".

In the following, a test procedure together with the early fail condition and the early pass condition is derived
30 with mathematical methods.

The ER test procedure is an iterative process: The number
ns of delays TD and the number of errors ne are accumulated from the beginning of the test to calculate a
35 preliminary ER (Error Ratio). Time delays TD up to the next error are used together with the past time delays TD to calculate the next preliminary ER. Every preliminary ER is tested against an early pass or early fail bound. The formulas to derive an early pass or early fail condition

contain a single step wrong decision probability D . This is valid only for one specific preliminary ER stage. However, it is meaningful to have a wrong decision probability F for the entire test ($F > D$). The application
5 proposes to use a single step wrong decision probability D , which is smaller than the probability for the entire test and proposes a method to derive D from a given F .

Due to the nature of the test, namely discrete error
10 events, the early stop conditions are declared not valid, when a fractional number of bad time delays < 1 are used to calculate the early stop limits. The application contains proposals, how to conduct the test at this undefined areas. The proposals are conservative (not
15 risky). A DUT on the ER-limit does not achieve any early stop condition. The application proposes to stop the test unconditioned at a specific number K (for example 200) bad time delays. As this proposal contains different measurement qualities, an additional proposal is made to
20 overcome this drawback.

Based on a measurement at a preliminary ER stage pERs, a confidence range CR around this measurement is derived. It has the property that with high probability the real
25 result can be found in this range.

The confidence range CR is compared with the specified ER limit. From the result a diagram is derived containing the early fail and the early pass condition.
30

With a finite number of delays n_s , the final Error Ratio ER cannot be determined exactly. Applying a finite number of delays n_s , a number of bad time delays n_e is measured. $n_e/n_s = er$ is the preliminary Error Ratio.
35

In a test at a preliminary ER stage pERs a finite number n_s of measured time delays TD is applied, and a number of bad time delays n_e is measured. n_e is connected with a certain differential probability in the Poisson

distribution. The probability and the position in the distribution conducting just one test at a preliminary ER stage pERs is not known.

- 5 Repeating this test infinite times, applying repeatedly the same ns, the complete Poisson distribution is obtained. The average number (mean value) of bad time delays is NE. NE/ns is the real ER at the DUT. The Poisson distribution has the variable ne and is characterised by
- 10 the parameter NE, the average or mean value. Real probabilities to find ne between two limits are calculated by integrating between such limits. The width of the Poisson distribution increases proportional to \sqrt{NE} , that means, it increases absolutely, but decreases
- 15 relatively.

In a test at a preliminary ER stage pERs ns time delays TD are applied and ne bad time delays are measured. The result can be a member of different Poisson distributions

20 each characterized by another parameter NE. Two of them are given as follows:

The worst possible distribution NE_{high} , containing the measured ne with the probability D_1 , is given by the

25 inverse cumulative operation

$$D_1 = \int_0^{ne} PD_{high}(ni, NE_{high}) dni \quad (1)$$

In the example D_1 is 0.002=0.2% PD_{high} is the wanted

30 Poisson distribution with the variable ni. ne is the measured number of bad time delays.

The best possible distributions NE_{low} , containing the measured ne with the probability D_2 is given by the

35 inverse cumulative operation

$$D_2 = \int_{ne}^{\infty} PD_{low}(ni, NE_{low}) dni \quad (2)$$

In the example D_2 is equal D_1 and it is $D = D_1 = D_2 = 0.002 = 0.2\%$.

5

To illustrate the meaning of the range between NE_{low} and NE_{high} refer to Fig. 3. Fig. 3 shows the likelihood density PD as a function of the measured number of bad time delays ne . In the example, the actual detected number of bad time delays ne within the measured number ns of time delays TD is 10. The likelihood distribution of the bad time delays is not known. The worst possible likelihood distribution PD_{high} under all possible likelihood distributions as well as the best possible likelihood distribution PD_{low} under all possible likelihood distributions are shown. The worst possible likelihood distribution PD_{high} is characterized in that the integral from 0 to $ne = 10$ gives a total probability of $D_1 = 0.002$. The best possible likelihood distribution PD_{low} is characterized in that the integral from $ne = 10$ to ∞ gives a total probability of $D_2 = 0.002$. In the preferred embodiment D_1 is equal to D_2 , i.e. $D_1 = D_2 = 0.002 = 0.2\%$. After having obtained the likelihood distribution PD_{high} and PD_{low} from formulas (1) and (2), the average values or mean values NE_{high} for the likelihood distribution PD_{high} and NE_{low} for the likelihood distribution PD_{low} can be obtained. The range between the mean values NE_{low} and NE_{high} is the confidence range CR indicated in Fig. 3.

30

In the case the measured value ne is a rather untypical result (in the example just 0.2% probability) nevertheless the real result NE can still be found in this range, called confidence range CR .

35

The probabilities D_1 and D_2 in (1) and (2) can be independent, but preferable they are dependent and equal ($D = D_1 = D_2$).

For the Poisson distribution NE_{low} and NE_{high} are the result of the inverse cumulative operations above and are shown in formulas (3) and (4) respectively. The inputs are
 5 the number of bad time delays ne , measured in this test, and the probabilities D and $C = 1 - D$. The Output is NE , the parameter describing the average of the Poisson distribution.

10 The following example is illustrated in Fig. 3 ($D = D_1 = D_2$):

$$NE_{low} = \frac{qchisq(D, 2 \cdot ne)}{2} \quad (3)$$

15
$$NE_{high} = \frac{qchisq(C, 2 \cdot ne)}{2} \quad (4)$$

Example:

Number of bad time delays: $ne=10$
 20 Probability: $D=0.002$ $C=0.998$

Result:

$NE_{low} = 3.26$
 25 $NE_{high} = 22.98$

Interpretation:

Having measured $ne=10$ bad time delays in a test with preliminary ER stage pERs, then at a low probability
 30 $D=0.002$ the average number of bad time delays NE in this test is outside the range from 3.26 to 22.98 or with a high probability $C=0.998$ inside this range from 3.26 to 22.98.

35 Such as the width of the Poisson distribution, the confidence range CR increases proportional to $SQR(ne)$, that means, it increases absolutely, but decreases relatively.

If the entire confidence range CR, calculated from a single result ne, is found on the good side ($NE_{limit} > NE_{high}$) of the specified NE_{limit} we can state: With high probability C, the final result NE is better than NE_{limit} .
Whereby NE_{limit} is given by

$$NE_{limit} = ER_{limit} \cdot ns \quad (5)$$

and ER_{limit} is the Error Ratio allowable for the device and obtained by an ideal long test with an infinite high number ns of time delays TD.

If the entire confidence range CR, calculated from a single result ne, is found on the bad side ($NE_{limit} < NE_{low}$) of the specified NE_{limit} we can state: With high probability C, the final result NE is worse than the ER_{limit} .

With each new time delay TD a new test is considered, reusing all former results. With each new test the preliminary data for ns, ne and er is updated. For each new test the confidence range CR is calculated and checked against NE_{limit} .

Once the entire confidence range CR is found on the good side of the specified limit ($NE_{limit} > NE_{high}$), an early pass is allowed. Once the entire confidence range CR is found on the bad side of the specified limit ($NE_{limit} < NE_{low}$) an early fail is allowed. If the confidence range CR is found on both sides of the specified limit ($NE_{low} < NE_{limit} < NE_{high}$), it is evident neither to pass nor to fail the DUT early.

Fig. 3 illustrates the above conditions. Of course, NE_{limit} is a fixed value not altering during the test, but NE_{low} and NE_{high} as well as the confidence range CR are altering during the test. For reasons of illustration, however, the three possibilities of the possible positions of the

confidence range CR with respect to the constant NE_{limit} are drawn for the same example in Fig. 3.

The above can be described by the following formulas:

5

The current number ns of time delays TD is calculated from the preliminary Error Ratio er and the preliminary number of bad time delays ne

$$10 \quad er = ne/ns \quad (6)$$

The specified ER expressed with number of time delays ns and number of bad time delays NE is

$$15 \quad ER_{limit} = NE_{limit} / ns \quad (7)$$

for abbreviation in the formula:

$$20 \quad \begin{aligned} er_{norm} &= er/ER_{limit} = ne/ NE_{limit} \\ &(\text{normalised } er) \end{aligned} \quad (8)$$

Early pass stipulates:

$$25 \quad NE_{high} < NE_{limit} \quad (9)$$

Early fail stipulates:

$$NE_{low} > NE_{limit} \quad (10)$$

30 Formula for the early pass limit:

$$er_{norm} = \frac{ne}{NE_{high}} \quad (11)$$

35 This is the lower curve ($ernorm_{pass}(ne, C)$) in Fig. 4, which shows er_{norm} as a function of ne .

Formula for the early fail limit:

$$er_{norm} = \frac{ne}{NE_{low}} \quad (12)$$

This is the upper curve ($er_{norm_{fail}}(ne, D)$) in Fig. 4.

5

As the early pass limit is not defined for $ne = 0$ (normally the case at the very beginning of the test for a good DUT), an artificial bad time delay event with the first measured time delay TD can be introduced. When the first real bad time delay event occurs, the artificial bad time delay is replaced by this real one. This gives the shortest possible measurement time for an ideal good DUT. For example $ns=50$ for $ER_{limit} = 0.1$ and probability $D = D_1 = D_2 = 0.2\%$. (6 minutes 20s)

15

As the early fail limit uses $NE_{low} < 1$ for small $ne < k$ (in the example below $k=5$) due to a decision problem at a fractional error, the early fail limit at $ne=k$ is extended with a vertical line upwards. This ensures that a very bad DUT hits the early fail limit in any case after a few delays, approx. 10 in the example (1 minute 20s). In other words, the test is not stopped as long as ne is smaller than k .

25 With each new time delay TD a new test is considered, reusing all former results. With each new test the preliminary data for ns , ne and er and er_{norm} are updated and a er_{norm}/ne coordinate is entered into the er_{norm} -diagram. This is shown in Fig. 5. Once the trajectory crosses the early fail limit ($er_{norm}(ne, D)$) or the early pass limit ($er_{norm}(ne, C)$) the test may be stopped and the conclusion of early fail or early pass may be drawn based on this instant.

35 Fig. 5 shows the curves for early fail and early pass. er_{norm} is shown as a function of the number of bad time delays ne . For the simple example demonstrated in Fig. 5, it is $ER_{limit} = 0.2 = 1/5$ and the final Error Ratio $ER =$

0.25 (1/4). The test starts with the first time delay TD, for which no error (bad time delay) is detected. For the second time delay TD, a first error (bad time delay) is detected and the preliminary Error Ratio $er = ne/ns = 1/2$ and $er_{norm} = er/ER_{limit}$ becomes $1/2 : 1/5 = 5/2$. er_{norm} after the second time delay TD is marked with a cross a in Fig. 5. For the third, fourth and fifth time delay TD, no further error (bad time delay) occurs and er_{norm} subsequently becomes 5/3, 5/4 and 5/5, respectively, which is marked with the crosses b, c and d in Fig. 5, respectively. The sixth time delay TD brings a new error (bad time delay) and ne becomes 2. Consequently, $er = ne/ns$ becomes 2/6 and er_{norm} becomes 10/6. This situation is marked with cross e in Fig. 5. For the seventh, eighth and ninth delay TD, no further error (bad time delay) occurs and the situation after the seventh, eighth and ninth delay TD is marked with crosses f, g, h in Fig. 5, respectively. The tenth time delay TD brings a third error (bad time delay). Consequently, er becomes 3/10 and er_{norm} becomes 15/10. This situation is marked with cross i in Fig. 5. As can be seen from Fig. 5, the trajectory is between the early fail curve and the early pass curve at the beginning of the test, but converges to a line Z, which crosses the early fail curve after about forty errors. After forty errors, it can thus be decided that the tested DUT early fails the test.

The calculations above give optimal results in terms of test-time for small error ratios /high success ratios. If this is not the case, the approach with the Poisson distribution can be used nevertheless. For any error ratio/success ratio the binomial distribution is valid. The Poisson distribution is wider than the equivalent binomial distribution, hence the confidence range is wider, hence the space between the early pass and the early fail limit is wider, hence the result is not optimal in terms of test-time, but conservative in the following sense: the test-time is longer than optimal and the wrong decision risk is lower than predefined.

To get optimal results also for higher error ratios/lower success ratios the binomial distribution must be applied as follows:

- 5 The binomial distribution has the following form:

$dbinom(ne, ns, ER)$

with

ne: number of errors (bad time delays) detected,

ns, number of time delays TD, measured,

- 10 $(ne/ns = \text{preliminary error ratio } er)$

ER: the mean value of the distribution is the error ratio

The worst possible distribution with respect to a measured $er=ne/ns$ is defined by the inverse cumulative operation:

15

$$D = \int_0^{ne} dbinom(ni, ns, ER_{high}) dni \quad (13)$$

with

D: wrong decision risk

- 20 ni: integration variable

ne: number of errors (bad time delays) currently detected, simultaneously integration limit

ns: number of time delays TD, currently measured

- 25 ER_{high}: The parameter to be tuned, to make the integral consistent. This is the wanted value.

The result has the form: $ER_{high} = qbinom(C, ne, ns)$

With $C=1-D$, the complementary wrong decision risk is equal to the confidence level.

- 30 The best possible distribution is defined by the inverse cumulative operation:

$$D = \int_{ne}^{ns} dbinom(ni, ns, ER_{low}) dni \quad (14)$$

- 35 with

ne: number of errors (bad time delays) currently detected,
simultaneously integration limit

ns: number of time delays TD, currently measured,
simultaneously integration limit

- 5 ER_{low}: The parameter to be tuned, to make the integral
consistent. This is the wanted value.

The result has the form: $ER_{low} = qbinom(D, ne, ns)$

- 10 Another inverse cumulative inversion, than the one above,
is frequently offered with standard mathematical software
tools. The form is similar: $qbinom(D, \text{at least one other}$
parameters than above).

- The confidence range is the range between ER_{high} and ER_{low}.
15 The normalized er_{norm} is the measured (preliminary)
 $er = ne/ns$ divided by ER limit

Early pass stipulates:

- 20 $ER_{high} < ER_{limit}$

Early fail stipulates:

$ER_{low} > ER_{limit}$

- 25 In analogy to the Poisson approach, the early fail limit
is

$$er_{norm\ fail} = er/ER_{low} \quad (15)$$

- 30 The early pass limit is

$$er_{norm\ pass} = er/ER_{high} \quad (16)$$

- The following consideration is valid for both, the Poisson
and the binomial approach:
35 If no early stop occurs the ER test may be stopped, after
the following condition is valid:

$$ne \geq K \quad (17)$$

and the DUT shall be passed, if n_s is sufficiently high. K is a maximum number of errors (bad time delays). For example K can be 200.

- 5 If the trajectory neither crosses the early fail curve nor the early pass curve after K (for example 200) errors (bad time delays) have occurred, the DUT can be finally passed. If the DUT, however, is rather good or rather bad, the tests can be stopped much earlier, long before the $K = 200$ errors (bad time delays) have occurred. This significantly shortens the total test time.

15 In the above embodiment early fail means: a DUT, which survived to certain preliminary ER stage, is failed and a probability of 0.2 % that it is actually better than the ER-limit is accepted. Further early pass means: the DUT, which survived to certain preliminary ER stage, is passed and a probability of 0.2 % that it is actually worse than the ER-limit is accepted. If the test is stopped at 200 errors (bad time delays) the DUT is passed without any early fail or early pass condition arbitrarily. It can cross the vertical 200 error line in Fig. 4 at different heights.

25 Each height is connected with a certain statistical interpretation: The probability to have a DUT better (worse) than the ER-limit is indicated in Fig. 7. The vertical in Fig. 7 shows the position in Fig. 4 at the end of the test. The horizontal line in Fig. 7 shows the respective probability. Hence such a test has different qualities, a better one at the early pass fail limits, and a worse and variable one at the unconditioned stop at 200 errors

35 The following embodiment overcomes this drawback and additionally accelerates the test. This is done by a meaningful redefinition of the early pass limit maintaining the early fail limit. Early pass means now: A DUT, which survived to certain preliminary ER stage, is

passed and a probability of 0.2 % that it is actually worse than M times the specified ER-limit ($M > 1$) is accepted. This is a bad DUT ER-limit. This shifts the early pass limit upwards in Fig. 4 as shown in Fig. 6.

5 $er_{norm\ pass} (ne, C)$ in Fig. 4 becomes $er_{normbad\ pass} (ne, C)$ in Fig. 6. $er_{norm\ fail} (ne, D)$ remains unchanged. Now it is

$$NE_{limit,M} = ER_{limit} \cdot M \cdot ns \quad \text{for Poisson}$$

$$ER_{limit,M} = ER_{limit} \cdot M \quad \text{for Binomial}$$

10 and an early pass is allowed, if

$$NE_{limit} \cdot M = NE_{limit,M} > NE_{high}. \quad (\text{Poisson})$$

$$ER_{limit} \cdot M = ER_{limit,M} > ER_{high} \quad (\text{binomial}) \quad (18)$$

There are three high level input parameters for the test:

- 15 - Specified error limit e.g. 10%
 - Probability to make a wrong decision in the entire test (proposal: $F = 3\%$)
 - Definition of a bad DUT ($ER_{limit} \cdot M$)

20 From this the working parameters are derived:

D, the wrong decision risk for a single test step, is derived from F, the wrong decision risk for the entire test. The method to do this is described below.

25 With D the early pass and fail limits can be calculated according to formula 11, 12, 15, 16.

The intersection co-ordinates of the early pass and the early fail limit are the maximum number of errors (bad time delay) and the error ratio test limit TL (specified error limit $< TL < M \cdot$ specified error limit).

30

There are infinite possibilities to overcome the above mentioned drawback with different test qualities.

35

In the example above the open end between the pass fail-limits was originally declared pass, and time was saved by multiplying the early pass limit with M ($M > 1$), shifting it upwards such that the early fail and the early pass curve

intersect at 200 errors (200 bad time delays: example from above). Thus only a DUT, bad with high probability, is failed (customer risk).

- 5 The complementary method is: The open end between the early pass fail limits is declared fail, and time is saved by multiplying the early fail limit with m ($0 < m < 1$), shifting it downwards, such that the early fail and the early pass curve intersect at 200 errors (200 bad time
10 delays: example from above). Thus only a DUT, good with high probability, is passed (manufacturer risk).

- The compromise method is: The open end between the early pass fail limits is partitioned in any ratio: the upper
15 part is declared fail and the lower part is declared pass. Time is saved by multiplying the early fail limit with m ($0 < m < 1$) and such shifting it downwards and by multiplying the early pass limit with M ($M > 1$) and such shifting it upwards. So the early fail and the early pass curve
20 intersect at 200 errors (200 bad time delays: example from above).

- With given D_1 and D_2 the early fail curve and the early pass curves in Fig. 5 and Fig. 4 or Fig. 6 can be
25 calculated before the test is started. During the test only $er_{norm} = ne/NE_{limit}$ (for binomial: $er_{norm} = er/ER_{limit}$) has to be calculated and to be compared with the early pass limit and the early fail limit as explained with respect to Fig. 5 and Fig. 6. Thus, no intensive
30 calculation has to be done during the test.

- D_1 and D_2 above describe the single step wrong decision probability for a DUT, which survived without early stop to a certain preliminary ER stage. For a real test it is
35 desirable to define in advance a wrong decision probability F for the entire test. For every step a fraction of DUTs leave the statistical totality due to single step wrong decision probability D . This accumulates to an amount $F > D$.

It is proposed to derive D from F by the following method:

Based on statistical independence, a large ensemble of
5 DUTs with ER on the ER-limit (limit-DUT) and with M*ER
(bad DUT) is simulated and the entire test for the limit-
DUT ensemble and the bad DUT ensemble is run by simulation
against the early pass and early fail bound, with a free
D-parameter (D_1 and D_2). The simulation will show, that a
10 certain fraction F ($D < F < 1$) of the large ensemble falsely
fails (limit-DUT) or falsely passes (bad DUT). This
represents the wrong decision probability F for the entire
test. D is tuned such that F corresponds to the predefined
wrong decision probability.

15 The initially mentioned wrong decision risk 3%
corresponds, as a result of the simulation, the single
step wrong decision risk of 0.2% applied in the formulas
and figures.

20 In retrospect the distributions are not any more exact
applying this simulation: Once a DUT leaves the
statistical totality due to early fail or early pass, the
subsequent distribution is changed. In so far the
25 initially selected distribution has its intended validity
only at the beginning of the test. Due to early passes and
early fails the distribution is truncated more and more
towards the end of the test. These truncated distributions
cannot be handled anymore analytically, and so it is
30 impossible to derive D from F analytically. It is possible
to derive each subsequent distribution by numerical
methods. Then it is possible to derive D from F
numerically. In the above proposed process, just a unique,
the initial distribution is applied as approximation
35 instead of a set of correct distributions. With a set of
correct distributions it is possible to calculate D from F
by:

$$F = 1 - (1 - D)^n \quad (19)$$

Generally it can be assumed:

$$F > D \geq 1 - (1 - F)^{1/ne} \quad (20)$$

5

This formula can serve as an infimum for D, if the method is based on simulation using the initial distribution.